

# THE IMPACT OF WAVE OF RESPONSE IN A LAW SCHOOL SURVEY ABOUT THE BELIEVABILITY OF WITNESSES

Thomas R. Belin, UCLA; Sandy Feinland, Turer and Feinland; Steven Winick, Shepherd Mullin Richter Hampton  
Thomas R. Belin, UCLA Dept. of Biostatistics, Los Angeles, CA, USA, 90095-1772

**Key words: nonresponse weights, exchangeability, ignorable nonresponse**

## Abstract

Members of a particular first-year section of the Class of 1990 at an eastern U.S. law school were exposed to a lecture where a woman introduced as a rape victim was later revealed to be an actress who was not a rape victim. A survey was carried out to assess whether this exposure had any long-term effect on attitudes toward the believability of witnesses. Graduates of this class were sent a mailed survey asking their opinions about various well-known public events (e.g., the Clarence Thomas-Anita Hill hearings) as well as their opinions about the percentage of criminal and civil trials in which at least some witness lies under oath. Because there was substantial nonresponse, wave of response was recorded, with three efforts made to contact each potential respondent. This paper compares alternative approaches to incorporating information about wave of response into comparisons between students who had been exposed versus unexposed to the particular lecture of interest.

## 1. Introduction

The study underlying the present work sought to examine the long-term effects (measured in 1998 and 1999) of one criminal law section at an Eastern U.S. law school (1987-1988 academic year) being exposed to a lecture where a woman was introduced as a rape victim and gave a tearful account of experience, took questions from students, then was revealed to be an actress. Of particular interest was whether such exposure would have long-term effects on attitudes about the believability of witnesses.

## 2. Methods

A survey was developed with questions about widely reported cases and news stories (specifically the O.J. Simpson trials, Clarence Thomas confirmation hearings, Bill Clinton inhaled marijuana smoke, William Kennedy Smith rape trial, Al Gore campaign funding solicitations, Marv Albert assault trial, Paula Jones lawsuit, and Monica Lewinsky's relationship with Bill Clinton). The division of students at the law school into sections presented us with a natural experiment for comparing exposed and unexposed students, since the actress/rape lecture occurred in only one section.

We obtained address information on students from alumni information. After excluding graduates with non-U.S. addresses, which was done to reduce survey costs, 524 students remained eligible. The first wave of surveys was mailed in late 1998, with follow-up of initial nonrespondents in early 1999. This was during the period that the impeachment and trial of Bill Clinton was taking place.

Study plans called for follow-up of initial non-response. Here, Wave 1 refers to responses to the original survey, which was mailed in December 1998; Wave 2 refers to responses after a reminder postcard was sent in February 1999 to encourage response to original survey; and Wave 3 refers to responses to a new copy of the survey that was sent in April 1999 to remaining non-respondents who had not refused.

We obtained 159 responses across the three waves. Respondent characteristics are summarized in Table 1 (categorical variables) and Table 2 (continuous variables). Subjects were asked on the survey to name their professors for each of their core first-year classes (Civil Procedure, Contracts, Criminal Law, Torts) so as not to draw attention to our interest in which Criminal Law section they were in. Classification into exposed versus unexposed was based on whether they named the professor who taught the section exposed to the actress/rape lecture. Study outcomes are described and summarized in Table 3. Likert-scaled variables were associated with values so that high scores corresponded to less sympathy for the accused.

Part of our interest here is to assess whether it is important to control for wave of response. We carried out 27 ANOVA's and 6 chi-square tests on characteristics and outcomes across waves. We found four significant results at  $p < 0.05$ . A binomial calculation yields  $\Pr(4 \text{ or more significant out of } 33) = 0.081$ , which does not provide strong evidence of a pattern. Results that were significant included the percentage of the time subjects believed some witness lies in a civil trial (means by wave: 48.9%, 63.3%, 61.0%;  $p = 0.039$ ), the average on a 7-point Likert scale from -3 = definitely not to 3 = definitely that people believed O.J. Simpson should have been held liable in his civil trial (means by wave: 1.51, 0.95, 0.82;  $p = 0.017$ ), and voter registration (percent of respondents who self-identified by wave as Democrats/Republicans/Independents/Other: 63/14/16/7, 40/25/30/5, 53/0/47/0,  $p = 0.037$ ). Also significant was the subjects' estimated difference between the percentage of civil trials in

which some witness lies and the percentage of criminal trials in which some witness lies (mean difference by wave: -10.5, -6.8, 4.3;  $p = 0.012$ ).

One of the only design variables available on all subjects was current zip code, which we categorized by first digit of zip code into (0, 1, 2, 3-4, 5-6, 7-8, 9). The response rate was lower for subjects with first digit of zip code = 1 (24/128 = 18.8%) than for other zip-code groups, which had response rates ranging from 30.2% to 43.2% ( $p = 0.006$ ). Analysis of variance by zip-code group revealed some differences in subject characteristics and outcomes. The overall response rate of 159/524 = 30.3%, giving rise to concern about nonresponse bias.

Grouping the number of responses from Waves 2 and 3 together, there was not a disproportionate number of early versus later responses in any zip-code group ( $p=0.643$ ).

We considered three alternative ways to weight survey respondents: (1) Weight respondents equally in unweighted analysis, (2) Assume non-response ignorable given region of current residence (i.e., increase weight for respondents from regions with below-average response rates), and (3) Assume non-response ignorable given region of current residence after Wave 1 (i.e., view nonrespondents as exchangeable only with late responders). Our primary planned analysis was to compare the mean in the exposed group ( $n_1=14$ ) and the mean in the unexposed group ( $n_2=145$ ) on 15 Likert-scaled outcomes and 2 continuously-scaled outcomes. Based on the theory that individuals exposed to actress/rape lecture would be less sympathetic to Anita Hill, Paula Jones, the accuser of William Kennedy Smith, and the accuser of Marv Albert, differences were expected a priori on 12 outcomes, and no differences were expected a priori on the remaining 5 outcomes.

### 3. Results

We considered whether the approach to weighting respondents affected conclusions about statistical significance by classifying findings according to whether  $p > 0.10$ ,  $0.05 < p < 0.10$ , or  $p < 0.05$ . Of the 17 outcomes we evaluated, weighting methods (1) and (2) agreed on 14 findings with  $p > 0.10$  and 1 finding with  $p < 0.05$ . Weighting methods (1) and (3) agreed on only 12 analyses with  $p > 0.10$ , reflecting the greater sensitivity that arises when unobserved cases are assumed exchangeable with the small numbers of cases (ranging from 4 to 8) within zip-code groups who responded during the latter two survey waves.

The significant finding in common between the unweighted analysis and the analysis assuming nonresponse ignorable within region was that exposed subjects were less likely to believe Secret Service

agents should be compelled to testify about conversations they overheard while in close proximity to President. This was in the expected direction based on theory that exposed would be more sympathetic to the accused (Clinton) than to the accuser (Paula Jones) in the underlying sexual harassment claim that triggered the Lewinsky investigation.

The analysis assuming non-response was ignorable within region (Approach 2) produced two findings of  $0.05 < p < 0.10$  that had  $p > 0.10$  in the unweighted analysis (Approach 1). First, exposed subjects were seen using Approach (2) to be less likely to believe Bill Clinton inhaled marijuana smoke. No difference had been expected a priori on this question. The analysis using weighting approach (3) found  $p < 0.05$  on this question. Second, using Approach (2), exposed subjects were less likely to believe Clarence Thomas ever said to Anita Hill that she would have a perfect case against him if she had a witness. The preamble to the question noted that Hill testified he said this, Thomas denied it, an administrative law judge friend of Hill's recalled Hill describing such a statement by Thomas years earlier, and a senator accused Anita Hill of perjury in her Senate testimony. The difference was in the a priori expected direction, although the effect was seen to have  $p > 0.10$  using Approach (3).

Weighting approach (3) featured two findings with  $p < 0.05$  that were contrary to the expected direction. These were that the exposed subjects were more likely to believe William Kennedy Smith was guilty of rape and that the exposed were less likely to believe Bill Clinton was "legally accurate" in his testimony in the Paula Jones lawsuit about his relationship with Monica Lewinsky. Both findings had  $p > 0.10$  using the other weighting approaches.

Pooling evidence across questions, we found that on questions where an exposure effect was anticipated, the direction of effect-size estimates agreed with the a priori expected direction 9 of 12 times using weighting approach (1), 8 of 12 times using weighting approach (2), and 7 of 12 times using weighting approach (3). Using Approach (1) and treating separate effect-size estimates as input to t-tests, a two-sample t-test comparing the 12 effect-size estimates where a difference was anticipated to the 5 where a difference was not anticipated produced a one-tailed p-value of 0.0888 and a two-tailed p-value of 0.1775. A one-sample t-test comparing the 12 effect-size estimates where a difference was anticipated to 0 produced a one-tailed p-value of 0.0544 and a two-tailed p-value of 0.1088.

We also carried out regression analyses for each of the outcomes listed in Table 3 (results not shown due to space limitations). Analyses controlled for exposure to the actress/rape lecture, gender, undergraduate major

area, party registration, an indicator for having clerked with a judge, opinion of legal education, personal outlook categories as described in Table 1 regarding believing there is such a thing as “truth”, and percentage of professional activities in each of the areas listed in Table 2. Due to the number of predictor variables, each outcome was fit at first to a model controlling for the various professional areas listed in Table 2 along with exposure to the target lecture, with a separate model relating each outcome to exposure and all of the other predictors. Effects with  $p < 0.20$  were then combined into an overall model. These models were supplemented with forward and backward stepwise regression procedures to search for other combinations of variables that might have been missed using the other variable selection procedure.

Of primary interest was to identify “consensus” predictors that seemed important no matter what other variables were included in the model. Again, exposure to the actress/rape lecture had  $p < 0.05$  only for the models fit to the outcome summarizing whether subjects believed Secret Service agents should be compelled to testify about conversations overheard in proximity to the President. In addition, in the models for the outcome summarizing whether job-related harm should be required for the defendant in a sexual harassment suit to be liable, exposure to the actress/rape lecture consistently had a negative coefficient with  $0.05 < p < 0.10$ . This was also in the expected direction, since higher values corresponded to belief that job-related harm should not be required to establish liability, a position supportive of Paula Jones and less sympathetic to the accused, Bill Clinton.

Some other variables showed up as significant predictors more often. For example, being registered Republican (vs. Democrat) had  $p < 0.05$  in 8 analyses and  $0.05 < p < 0.10$  in 2 more; opinion of legal education had  $p < 0.05$  in 3 analyses and  $0.05 < p < 0.10$  in 6 more; gender had  $p < 0.05$  in 2 analyses and  $0.05 < p < 0.10$  in 3 more; and percentage of time working in plaintiff civil litigation had  $p < 0.05$  in 2 analyses on  $0.05 < p < 0.10$  in two more. On the other hand, having a humanities vs. social science undergraduate major had  $p < 0.05$  in one analysis and  $0.05 < p < 0.10$  in one other, and having a science vs. social science undergraduate major never had  $p < 0.05$  and twice had  $0.05 < p < 0.10$ .

The sampling distribution of the number of significant variables per analysis is extremely complicated given the hybrid model selection procedure used, but assuming that of the 21 predictors one might expect  $21 \times 0.05 = 1.05$  variables per analysis to emerge in the category  $p < 0.05$  and the category  $0.05 < p < 0.10$  purely by chance under the null hypothesis of no effect of any predictors, then across 17 analyses one

would expect about 18 predictors to fall into these categories. We observed 26 findings with  $p < 0.05$  across all 17 analyses and 33 findings with  $0.05 < p < 0.10$ , suggesting that the findings are not all due to chance alone.

#### 4. Discussion

The roughly 30% response rate led to a fair degree of sensitivity in study findings. Substantial concerns arise in this analysis over multiple tests, as a certain number of significant results are to be expected by chance even in the absence of an exposure effect.

On the other hand, because the exposed section was a small section, the study also had limited statistical power to detect exposure effects. The analyses pooling findings across various study outcomes come close to significance, especially under the seemingly reasonable approach of applying a one-tailed test given the clear directional hypothesis underlying the study.

While we can conclude that other factors influence people’s attitudes on issues to a greater extent, the possibility of small effects due to the target lecture remains plausible. We believe that the merits of techniques such as introducing a classroom speaker as having a background that the person really doesn’t have, which some might consider deceptive, deserve to be discussed further. Although the findings of the present study are ambiguous, we nevertheless think that they may be valuable in such discussions.

**Table 1. Categorical characteristics of respondents**

Variable	n	% of responders
Gender		
Male	58	36.5
Female	101	63.5
Undergraduate major		
Humanities	48	31.0
Social science	86	55.5
Science	21	13.5
Clerked for a judge?		
Yes	106	67.1
No	52	32.9
Personal outlook*		
No such thing as the "truth"	13	8.4
There is "truth"	112	72.7
Neither of these	29	18.8

Variable	n	% of responders
Exposed to actress/rape lecture?		
Yes	14	8.8
No	145	91.2
Voting registration		
Democrat	91	59.5
Republican	22	14.4
Independent	31	20.3
Other/not registered to vote	9	5.9
Overall opinion of law school education		
Excellent	64	43.0
Good	69	46.3
Fair	15	10.1
Poor	1	0.7

\* Exact question wording: "Which of the following best describes your personal outlook?" Choices: (i) I believe that there is no such thing as the "truth", since the same events can be perceived differently by different people, (ii) I believe that there is such a thing as the "truth", even though it may not always be discernible, (iii) Neither of these.

**Table 2. Continuous characteristics of respondents**

Professional activities	n	Mean	SD	Proportion > 0% time	Proportion ≥ 50% time
% time working in criminal prosecution	154	1.23	4.93	.104	.000
% time working in criminal defense	154	4.45	14.17	.260	.032
% time in plaintiff civil litigation work	157	15.86	23.32	.592	.108
% time in defense civil litigation work	156	26.92	31.12	.724	.282
% time in corporate non-litigation work	155	28.74	37.14	.639	.265
% time working as law professor	152	3.59	13.76	.132	.033
% time in other work as an attorney	151	12.95	22.27	.470	.126
% time not working as attorney	149	8.86	22.59	.201	.087

**Table 3. Descriptive summaries of outcome variables**

Outcome variable	Scoring	Mean (SD)	Distribution of scores
Simpson guilty in criminal trial?	3 = definitely, 2 = almost certainly, 1 = probably, 0 = not at all clear, -1 = probably not, -2 = almost certainly not, -3 = definitely not	1.03 (1.40)	Score: -3 -2 -1 0 1 2 3 Count: 1 6 17 32 33 46 23
Simpson liable in civil trial?	3 = definitely, 2 = almost certainly, 1 = probably, 0 = not at all clear, -1 = probably not, -2 = almost certainly not, -3 = definitely not	1.37 (1.18)	Score: -3 -2 -1 0 1 2 3 Count: 1 1 3 38 34 53 28
Simpson: double jeopardy?	1 = no, 0 = not sure, -1 = yes	0.91 (0.38)	Score: -1 0 1 Count: 5 4 148
Clarence Thomas: did he say Anita Hill would have had perfect case if she had had a witness?	3 = definitely, 2 = almost certainly, 1 = probably, 0 = not at all clear, -1 = probably not, -2 = almost certainly not, -3 = definitely not	0.53 (0.86)	Score: -3 -2 -1 0 1 2 3 Count: 1 0 8 74 53 15 3
Thomas: should he have been confirmed?	1 = no, 0 = not sure, -1 = yes	0.24 (0.88)	Score: -1 0 1 Count: 46 26 84
Clinton inhaled marijuana smoke?	3 = definitely, 2 = almost certainly, 1 = probably, 0 = not at all clear, -1 = probably not, -2 = almost certainly not, -3 = definitely not	1.45 (1.21)	Score: -3 -2 -1 0 1 2 3 Count: 1 1 10 17 44 53 31
William Kennedy Smith guilty of rape?	3 = definitely, 2 = almost certainly, 1 = probably, 0 = not at all clear, -1 = probably not, -2 = almost certainly not, -3 = definitely not	0.17 (0.89)	Score: -3 -2 -1 0 1 2 3 Count: 1 2 22 92 28 11 2
Al Gore: would appoint independent counsel re campaign funding solicitations?	3 = definitely, 2 = almost certainly, 1 = probably, 0 = not at all clear, -1 = probably not, -2 = almost certainly not, -3 = definitely not	-0.62 (1.50)	Score: -3 -2 -1 0 1 2 3 Count: 20 26 41 33 28 7 4
Marv Albert trial: rule against asking alleged victim re sexual history too restrictive?	3 = definitely not, 2 = almost certainly not, 1 = probably not, 0 = not at all clear, -1 = probably, -2 = almost certainly, -3 = definitely	0.13 (1.14)	Score: -3 -2 -1 0 1 2 3 Count: 7 4 14 91 27 11 5
Supreme Court giving ok to Paula Jones suit while Clinton in office a good decision?	2 = definitely, 1 = probably, 0 = not at all clear, -1 = probably not, -2 = definitely not	-0.63 (1.43)	Score: -2 -1 0 1 2 Count: 57 51 4 26 20

Outcome variable	Scoring	Mean (SD)	Distribution of scores
Paula Jones lawsuit: job-related harm required for sexual harassment liability?	2 = definitely not, 1 = probably not, 0 = not at all clear, -1 = probably, -2 = definitely	-0.50 (1.35)	Score: -2 -1 0 1 2 Count: 43 55 10 33 15
Clinton/Lewinsky: should perjury about relationships generally be prosecuted?	2 = definitely, 1 = probably, 0 = not at all clear, -1 = probably not, -2 = definitely not	-0.13 (1.27)	Score: -2 -1 0 1 2 Count: 24 51 19 49 15
Clinton/Lewinsky: his testimony in Jones suit "legally accurate"?	1 = no, 0 = not sure, -1 = yes	0.41 (0.79)	Score: -1 0 1 Count: 30 34 94
Clinton/Lewinsky: were their statements material to Paula Jones lawsuit?	2 = definitely, 1 = probably, 0 = not at all clear, -1 = probably not, -2 = definitely not	-0.47 (1.21)	Score: -2 -1 0 1 2 Count: 34 60 20 35 9
Clinton/Lewinsky: should Secret Service agents be compelled to testify about President?	2 = definitely, 1 = probably, 0 = not at all clear, -1 = probably not, -2 = definitely not	-0.69 (1.22)	Score: -2 -1 0 1 2 Count: 45 63 15 24 10
In what percent of criminal trials does some witness lie under oath?	Recorded as percentage between 0 and 100	60.36 (26.11)	Value: 0 1-9 10-19 20-29 30-39 40-49 50-59 60-69 70-79 80-89 90-99 100 Count: 0 2 2 14 13 5 29 8 16 18 18 10
In what percent of civil trials does some witness lie under oath?	Recorded as percentage between 0 and 100	51.72 (26.52)	Value: 0 1-9 10-19 20-29 30-39 40-49 50-59 60-69 70-79 80-89 90-99 100 Count: 0 1 9 20 15 6 35 10 15 8 6 12